

令和 5 年 6 月 20 日現在

機関番号：53901

研究種目：挑戦的研究（萌芽）

研究期間：2018～2022

課題番号：18K18517

研究課題名（和文）手話認識システムを利用した手手辞典の開発と手話による百科事典の提案

研究課題名（英文）Development of a sign language dictionary using a sign language recognition system and a proposal for an encyclopedia in sign language

研究代表者

木村 勉（KIMURA, Tsutomu）

豊田工業高等専門学校・情報工学科・教授

研究者番号：80225044

交付決定額（研究期間全体）：（直接経費） 4,000,000円

研究成果の概要（和文）：本研究では、手話認識機能を搭載したオンライン手話単語辞書システムを開発した。システムはブラウザから利用する。Webカメラを用いて調べたい手話を表現すると、それを認識して、候補単語の一覧を表示する。利用者が表示された単語をクリックすると、見本動画が再生し、確認することができる。現在は、175単語の認識が可能である。さらに、ここで開発した手話単語認識エンジンを応用して、手話学習システムの開発を行っている。これらは2023年度中に一般に公開する予定である。

研究成果の学術的意義や社会的意義

手話を学習するためには、手話サークルに通うなど、対面で習う必要がある。テレビなどでも学習できるが、自身の手話が相手に伝わるかどうかはわからない。英語学習であれば、音声認識を使ったシステムもあるが、手話にはそれがなかった。本研究は、手話学習者に大いに役立てると考えている。さらに、本研究で開発した手話認識用学習ネットワークは、単語だけでなく、手話文、指文字にも利用できる。今後、手話認識システムの開発にも応用できる。

研究成果の概要（英文）：We have developed an online sign language dictionary system with a sign language recognition function. The system is designed to be used through a web browser. When a user performs a sign language gesture using a webcam, the system recognizes it and displays a list of candidate words. When the user clicks on a word, a sample video is played back for confirmation. Currently, the system can recognize 175 words. Furthermore, we are currently developing a sign language learning system by incorporating the sign language word recognition engine developed in this research. These systems will be released to the public in FY2023.

研究分野：福祉工学

キーワード：手話認識 手話辞書 ディープラーニング 深層学習 手話学習

1. 研究開始当初の背景

我々はこれまで手話から日本語を検索することのできる「日本手話・日本語辞書システム」を開発してきた。この理由として、日本語から手話を調べるといふ日本語手話辞書が多く存在するのに対し、手話からそれに対応する日本語を引き当てる辞書(日本手話・日本語辞書)の数が少ないためである。図1に開発した日本手話・日本語辞書システムのGUIを示す。

この日本手話・日本語辞書システムは、手話表現において6つに分類した手話音素から調べたい手話に当てはまるものを選んでいき、目的の手話を検索するという仕組みである。しかしそのために手話の初心者には使いにくいシステムであった。また先天性聴覚障がい者の場合、多くは書記日本語に明るくない。そのため知らない手話を調べようとしても、GUIに書かれている選択肢の日本語の理解が難しいという問題があった。

これらの問題を解決する方法として、手話認識システムを導入する。利用者がカメラの前で手話単語を表現すれば、システムが手話を認識し、候補の一覧を画面に表示する。利用者は、その一覧の中から調べたい手話を探ることができるようにする。



図1 日本手話・日本語辞書システムのGUI

2. 研究の目的

本研究で開発する手話辞書システムの構成を図2に示す。本システムでは次の利用方法を考えている。1) ユーザがパソコンやスマートフォンのカメラの前で手話を表現する。2) システムがそれを認識し、複数の手話語彙候補の一覧をサムネイル動画として表示する。3) ユーザは一覧の中から手話語彙を選択する。4) その手話の意味を手話やイラスト・平易な日本語などで解説する。

さらにこの研究成果を基に手話による百科事典の提案も行う。これは、調べたい物事(書記日本語)の意味を手話などで解説するものである。現在ネット上では Wikipedia のように手軽に物事を調べることができるが、当然ながらその解説は書記日本語であり、書記日本語に明るくない人にとっては使いにくい。これを手話によって解説するのである。

過去の研究(手話による情報保障システム)においても、聾者には書記日本語では知らない言葉があってもわかりにくい、手話の方がわかり易いという評価を得ており、聾教育においても重要であると考えられる。

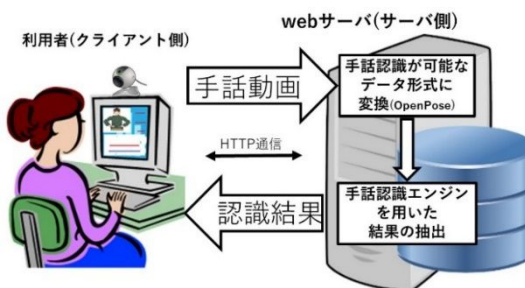


図2 システム概要図

3. 研究の方法

手話辞書システムを実現するために、本研究課題では以下のことを順次行う。

- 1) 語彙レベルの手話動画の収集

本研究で用いる学習用データの収集を行う。手話モデルに依頼し、手話技能検定の試験範囲を基準として語彙の動画を収集する。手始めに6級の試験範囲である101語について、それぞれ100個のデータ(計10,100個)を撮影し、アノテーションを行う。

これらの動画から姿勢推定ライブラリ(当初はOpenPose、後にMediaPipeに変更)を用いて関節などのキーポイントの座標を推定する。推定したキーポイントは動画の1フレームごとに1つのJSON形式ファイルとして出力される。つまり30fpsの動画の場合、1秒間で30個のファイルとなる。これを動画ごとに1つのCSV形式ファイルにまとめる。また動画によって解像度が異なると、キーポイントの座標系が異なってしまう。どのような解像度でも同じ座標系にするため、正規化を行って、統一した座標系に変換する。
- 2) ディープ・ラーニング手法を用いた手話認識エンジンの開発

機械学習のネットワークとしては自然言語処理で飛躍的な性能をもっているTransformerと画像認識などに用いられるCNNを組み合わせたConformerを適用させる。学習は、語彙の意味と座標データをペアとする教師あり学習とし、学習済みモデルを作成する。その学習済みモデルを用いて、座標データから手話を認識するシステムを構築する。
- 3) 手話辞書システムの開発

今回のシステムでは、GUI はできるだけシンプルな構成にする。利用者は手話の録画と送信、認識結果の一覧から、単語をクリックするだけの操作となる。そこで図3に示すような構成とする。

システムを起動すると PC に接続されたカメラと接続し、動画プレーヤに表示する。録画ボタンを押すと録画が開始される。録画した動画データを辞書システムにアップロードすることで認識が始まる。

認識が終わると、画面右側に確率の高い順に10個の単語が一覧として表示される。この単語をクリックすることで、見本の手話が再生される。利用者は、順次単語をクリックして再生することで、自分が表現し、調べたい手話の意味を確認することができる。

本システムは、最終的にはオンラインで利用できることを目指している。そこで GUI はブラウザ上で動作させることとし、HTML、CSS、JavaScript を用いて実装する。

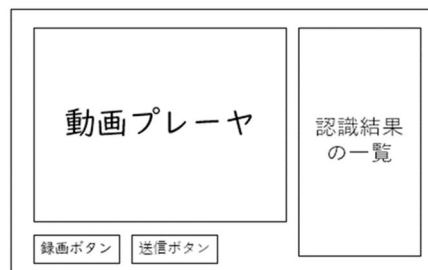


図3 GUIの基本設計

4. 研究成果

4.1 手話認識エンジンの実装

Conformer を用いて学習させた結果、単語ごとの認識率にばらつきがあるものの、約96.1%の認識率が得られた。また、学習させる語彙の種類を手話技能検定試験6級(101単語)に加え、5級の一部(78単語)を追加した。

4.2 手話辞書システムの実装

図4に起動時の画面を示す。中央に動画プレーヤ、右側に結果の一覧、下部に操作のボタンを配置している。利用者は、録画ボタンを押して、調べたい手話を表現する(図5)。録画したデータは一旦、ローカルストレージに保存する。録画した動画をサーバにアップするが、録画したデータ以外に、あらかじめ用意したデータをアップすることもできる。

サーバにアップすることで、認識が始まる。認識が終わると、図6のように画面右側に確率の高い順に10個の単語が提示される。利用者が単語名をクリックすると、図7のように見本の手話動画が再生され、調べたい手話かどうかを確認することができる。



図4 GUI初期画面



図5 検索したい手話の録画



図6 認識結果の提示



図7 見本手話の確認

なお、これらのシステムについては、下記のサイトで公開する予定である。

<https://jsldic.net>

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 Tsutomu Kimura, Kazuyuki Kanda	4. 巻 31
2. 論文標題 Sign Language Recognition through Machine Learning by a New Linguistic Framework	5. 発行年 2019年
3. 雑誌名 Technology and Disability	6. 最初と最後の頁 137-146
掲載論文のDOI（デジタルオブジェクト識別子） 10.3233/TAD-190008	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計10件（うち招待講演 0件 / うち国際学会 2件）

1. 発表者名 池田 康希, 木村 勉
2. 発表標題 ディープ・ラーニングを用いた手話認識に関する研究 ~ 2単語文と手話辞書構造を導入したConformerによる認識 ~
3. 学会等名 第121回 福祉情報工学研究会
4. 発表年 2023年

1. 発表者名 磯谷 光, 木村 勉, 神田 和幸
2. 発表標題 ディープ・ラーニングを用いた手話認識に関する研究 ~ CTCとConformerの比較 ~
3. 学会等名 第116回 福祉情報工学研究会
4. 発表年 2022年

1. 発表者名 磯谷 光, 木村 勉, 神田 和幸
2. 発表標題 ディープ・ラーニングを用いた手話認識に関する研究 ~ 手話文からの単語認識 ~
3. 学会等名 第111回 福祉情報工学研究会
4. 発表年 2021年

1. 発表者名 木村 勉, 鈴木 洸輔, 石濱 春, 吉田 あすか, 神田 和幸
2. 発表標題 手話認識機能を備えた手話辞書システムの開発
3. 学会等名 第111回 福祉情報工学研究会
4. 発表年 2021年

1. 発表者名 木村 勉, 神田 和幸
2. 発表標題 類似手話語彙の平面データによる光学的識別法(1) 特徴点の抽出と遷移の検証
3. 学会等名 第18回情報科学技術フォーラム
4. 発表年 2019年

1. 発表者名 神田 和幸, 木村 勉
2. 発表標題 類似手話語彙の平面データによる光学的識別法(2) 平面データと立面データの比較
3. 学会等名 第18回情報科学技術フォーラム
4. 発表年 2019年

1. 発表者名 Kimura Tsutomu, Kanda Kazuyuki
2. 発表標題 Sign Language Recognition through Machine Learning by a New Linguistic Framework
3. 学会等名 Association for the Advancement of Assistive Technology in Europe 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 高橋 佑汰, 木村 勉, 神田 和幸
2. 発表標題 機械学習を用いた手話認識に関する研究
3. 学会等名 福祉情報工学研究会
4. 発表年 2019年

1. 発表者名 神田 和幸, 木村 勉
2. 発表標題 描素の構造 ~ 新手話学の演繹的アプローチ ~
3. 学会等名 福祉情報工学研究会
4. 発表年 2019年

1. 発表者名 T. Kimura, Y. Takahashi, K. Kanda
2. 発表標題 Study on sign language recognition using machine learning
3. 学会等名 Association for the Advancement of Assistive Technology in Europe 2019 (国際学会)
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担 者	神田 和幸 (KANDA Kazuyuki) (70132123)	国立民族学博物館・人類基礎理論研究部・外来研究員 (64401)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------